PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
G01N 33/569, 33/564, 33/566, C07K 14/705

A1

(11) International Publication Number: WO 98/59244

(43) International Publication Date: 30 December 1998 (30.12.98)

(21) International Application Number: PCT/GB98/01801

(22) International Filing Date: 19 June 1998 (19.06.98)

(30) Priority Data:
9712892.0 20 June 1997 (20.06.97) GB

(71) Applicant *(for all designated States except US)*: ECLAGEN LIMITED [GB/GB]; Marischal College, Broad Street, Aberdeen AB9 1AS (GB).

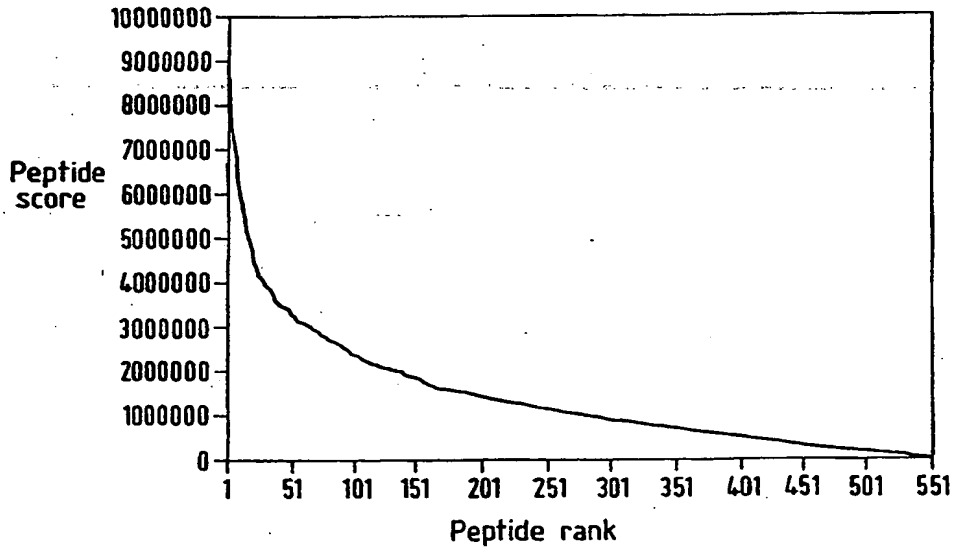
(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: FOTHERGILL, John [GB/GB]; The Granary, Shorehead AB39 2JY, Stonehaven (GB). KEMP, Graham [GB/GB]; Netherton, Sauchen AB59 7JP, Inverurie (GB). BROOKS, Tony [GB/GB]; 10 Sunnybank Place, Aberdeen AB24 3LA (GB). CARR, Frank [GB/GB]; Auris, 23 St. Machar Drive, Aberdeen AB24 3RY (GB).

(74) Agents: STEBBING, Peter, John, Hunter et al.; Ablett & Stebbing, 45 Lancaster Mews, Lancaster Gate, London W2 3QQ (GB).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, GW, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: IDENTIFICATION OF MHC BINDING PEPTIDES



(57) Abstract

The invention provides a method for the prediction of the binding affinity of a peptide to a major histocompatilibity (MHC) class II molecules comprising; 1) ascertaining the characteristics of a MHC molecule binding groove, 2) presenting a selected peptide to the MHC molecule and ascertaining a first conformation score for each pocket bound peptide side–chain, 3) amending the conformation of each pocket bound peptide side–chain and ascertaining a second conformation score, 4) repeating step 3 with alternative conformations of each peptide pocket bound side–chain, 5) choosing the highest conformation score for each pocket bound peptide side–chain in each binding groove pockets, herein known as "the pocket", and 6) combining the highest conformation score for each pocket and ascertaining a binding score for the complete peptide.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

IDENTIFICATION OF MHC BINDING PEPTIDES

The present invention relates to a new method for the prediction of peptides which bind to major histocompatibility (MHC) class II molecules and to molecules created or modified through the use of these methods.

The immune system of the mammalian organism principally comprises two arms, the cellular immune system and the humoral or antibody-associated immune system. The cellular immune system is centred around the activity of T cells. There are two major classes of T cells, cytotoxic T lymphocytes (CTLs) which attack cells displaying foreign antigen complexed with MHC class I molecules, and helper T cells which react to cells displaying foreign antigens in a complex with MHC class II molecules resulting in the secretion of cytokines which can activate B cells to produce antibody molecules.

Humans express six different MHC class I genes and six different MHC class II genes, which are located on three highly polymorphic loci. This leads to considerable allelic variation in MHC molecules. The MHC class I consist of a α-chain and a β_2-microglobulin, the α-chain is split into three domains α_1, α_2 and α_3. α_1 and α_2 form the MHC class I binding groove which contains pockets that bind the side chains and the amino and carboxy termini of any peptide present in the groove. The MHC class II molecules comprise an α-chain and a β-chain, it is the α_1 and β_1 domains which create the MHC class II binding groove. The MHC class II binding groove also contains pockets but it does not bind the end termini of the peptide. For this reason the peptides bound by the MHC class II molecule can be longer and of a more variable length. The typical length of peptides complexed with a MHC class I or a MHC class II molecule are 8-10 amino acids and 13-20 amino acids, respectively.

At present only three MHC class II structure are available but

it is believed that the backbone structure of all MHC class II alleles presently identified are similar to that of HLA-DR1. Structures of different alleles can be predicted by using homology modelling. This involves identifying the amino acid differences near the binding groove and using a computer to change the conformation of the side-chains to give favourable steric and electrostatic arrangements and to make the pockets as large as possible. The end result is a three dimensional structure of a MHC class II molecule, which can be used in various experiments.

The ability to predict the peptides in a protein which can bind to a given MHC molecule has great value especially for medical applications. It is known, for example, that in certain auto-immune diseases, T cells react with self-peptides presented by MHC class II molecules. It would be valuable to predict which peptides from auto-immune proteins are presented by MHC class II molecules in these diseases as well as to predict the binding of analogues of these peptides synthesised as potential antagonists for the presentation of self-peptides. In the selection of peptides for synthetic vaccines, the ability to predict MHC class II binding peptides would be advantageous. In addition, where heterologous proteins are developed as medicines or diagnostic imaging agents, it would be advantageous to predict potential MHC class II binding peptides in order to eliminate these from the heterologous proteins before administration to patients.

While studies of peptides complexed with MHC class I molecules have revealed conserved "anchor" residues at certain positions within the presented peptides, such studies with peptides complexed with MHC class II molecules have been less successful mainly because of the greater length variability of such peptides and the consequent difficulty in aligning their sequences.

Methods for accurately predicting the binding potential of

peptides have been restricted to MHC class I interaction with a peptide. In one method using three-dimensional structures of MHC class I molecules, peptide binding is ranked in ascending order according to the energy values determined. This method requires that the MHC structure be known, or that there is an obvious molecular model for the MHC structure. An identical method is said to be available for MHC class II but it does not consider the longer average length of the peptide and the open-ended peptide binding groove of MHC class II molecules. Neither does it use the best potential conformation of peptide amino acid side-chains and, therefore the binding energies calculated are only approximations.

Another drawback of using the same method for MHC class I and MHC class II peptide binding is that the binding of peptides to MHC class II is less dependant on strict allele-specific binding motifs than peptides binding to MHC class I. Individual amino acids in the peptide play a more significant role in MHC class II binding than MHC class I such that the conformation of amino acid side-chains is proportionally more important to the accuracy of binding analysis. Therefore, known methods do not provide a general method for analysing the binding of peptides to three-dimensional structures of MHC class II. There is thus a need for improved methods for predicting the MHC class II binding potential of peptides.

An object of this invention is to provide a method for accurately predicting the binding affinity of a peptide fragment binding to a MHC class II molecule.

Another object of this invention is to provide a computer conditioned to perform the task of predicting the binding affinity of a peptide fragment binding to a MHC class II molecule.

A yet further object of this invention is to provide a vaccine derived from the peptide fragment whose binding affinity to

MHC class II molecules has been determined.

Another object of this invention is to provide a pharmaceutical composition which comprises a peptide whose binding affinity to MHC class II molecules has been determined.

According to the first aspect of this invention, there is provided a method for the prediction of the binding affinity of a peptide and a major histocompatibility (MHC) class II molecules comprising;

1) ascertaining the characteristics of a MHC molecule binding groove,

2) presenting a selected peptide to the MHC molecule and ascertaining a first conformation score for each pocket bound peptide side-chain,

3) amending the conformation of each pocket bound peptide side-chain and ascertaining a second conformation score,

4) repeating step 3 with alternative conformations of each peptide pocket bound side-chain,

5) choosing the highest conformation score for each pocket bound peptide side-chain,

6) combining the highest conformation score for each pocket-bound peptide side-chain and then ascertaining a binding score for the peptide.

It is particularly desirable to then compile information on all peptide fragments in a protein and compare the binding scores. It is preferable if the conformation of the backbone of the peptide fragment is also altered and the conformation score and the binding score is then reassessed.

The method of this invention thus involves assessing a binding score for all possible candidate peptides by considering the predicted three-dimensional conformations and interactions between the MHC and the peptide in the complex. The computed score indicates the predicted binding affinity for the

particular peptide binding with the MHC allele and can be used to predict whether the peptides are likely to bind, or not.

Preferably, the conformation score for each pocket bound peptide side-chain is ascertained by considering at least one of the following parameters:

a) the steric overlap between the pocket bound peptide residue bound in the pocket and an atom forming the pocket; this is value B,

b) the number of hydrogen bonds which can be formed between the pocket bound peptide residue and an atom forming the pocket; this is value C,

c) the strength of electrostatic interactions between any polar atoms of the pocket bound peptide residue and any polar atoms forming the pocket; this is value D, and

d) the number of favourable contacts between the pocket bound peptide residue and the MHC residues forming one of the pockets; this is value E.

The conformation score for each peptide is computed based upon the predicted atomic interactions between each of the pocket bound peptide residues and MHC pockets. The geometric constraints imposed on the peptide by the shape of the MHC binding groove play an important part of the scoring function. Favourable packing arrangements between peptide and MHC side-chains are rewarded by the scoring function, whilst arrangements involving steric overlap are penalised. Alternative conformation are tried for MHC residues if an MHC residue overlaps with a peptide side chain.

If no preferable conformation can be found the MHC side-chain is returned to its original conformation. In the event of more than a pocket residue side-chain overlapping with a pocket bound peptide side chain, the pocket residue side chains are adjusted in order of overlap severity, with the pocket residue side-chain which has the most severe overlap being adjusted first.

In preferred embodiments the steric overlap between the pocket bound peptide residue and the atoms forming the pocket can not be greater than 0.35 Angstroms, otherwise the residue is deemed unable to fit in the pocket.

Conveniently a favourable contact occurs when an atom from an MHC residue and an atom from the peptide residue have their centres separated by no more than the sum of their radii plus 0.5 Angstroms and are not overlapping.

Preferably the values B to E are imported into a first equation to give a conformation score(Z). The first equation is $Z_n=(cK_2C)-(cK_3D)+(cK_4E)-(cK_1B)$, where cK_1 to cK_4 are constants and n is the number of the pocket.

The value of cK_1 is between 50 and 150. Preferably between 75 and 125.

The value of cK_2 is between 1000 and 2000. Preferably between 1250 and 1750.

The value of cK_3 is between 250 and 750. Preferably between 350 and 650.

The value of cK_4 is between 500 and 1500. Preferably between 750 and 1250.

Conveniently the Z_n value for a pocket is multiplied by a coefficient, L, depending on the pockets importance in binding, to give a second Z_n value. The value L is in the range of 0.001 to 5. Larger pockets are considered more important in determining which peptide can bind, compared with the other smaller pockets, so the scores contributed by each pocket are weighted in proportion to the amount of the peptide side-chain buried by the surface of the MHC molecule. When binding to MHC class II molecules, peptides have shown high similarity in the degree to which their side-chains are buried

by the MHC surface, despite having dissimilar sequences.

Preferably all the Z_n values are summed to give a value J. Value J is the overall contributing score of all the pockets for a certain conformation of the peptide fragment.

Conveniently the MHC residue is paired with the pocket-bound peptide residue if an atom from the MHC residue and an atom from the pocket-bound peptide residue have their centres separated by no more than the sum of their van der Waal radii plus one Angstrom.

In a preferred embodiment a value A_n is calculated by summing the pairwise interaction frequencies of paired residues. As for the Z_n value, preferably the value A_n for a pocket is multiplied by a coefficient, X, depending on the pockets importance in binding. Preferably X is between 0.001 and 5.

Conveniently the A_n value for the pockets are summed to give a value P.

In a preferred embodiment the binding score is ascertained by at least one of the following parameters

a) the number of groove-bound hydrophobic residues; this is value F,

b) the number of non groove-bound hydrophilic residues; this is value G,

c) the number of peptide residues deemed to fit within their respective binding pocket; this is value H.

Preferably values F, G, H, J and P are imported into a second equation to give a first binding score, Y.

Conveniently the second equation is $Y=J*F^2*(G*H+1)+P$.

However, in the alternative, the term He, which evaluates the hydrophobicity of the pocket bound peptide side chains using

a hydrophobicity scale disclosed in Janin et al [1979] Nature, 277 pg 491, can also be used to determine the Y value. Accordingly, $Y=(bK_2C)-(bK_3D)+(bK_4E)-(bK_1B)+(bK_5He)+P$. The scale used in Janin et al to measure hydrophobicity has a range from -1.8 for lysine to 0.9 for cysteine.

It is known that peptides having favourable hydrophobic/hydrophobic interactions with solvent and MHC atoms have a higher binding affinity. Accordingly, it is preferable to include the term He.

The value of bK_1 is between 1 and 10. Preferably between 1 and 5.

The value of bK_2 is between 20 and 60. Preferably between 30 and 50.

The value of bK_3 is between 300 and 900. Preferably between 450 and 750.

The value of bK_4 is between 1 and 20. Preferably between 5 and 15.

The value of bK_5 is in between 1 and 800. Conveniently between 100 and 600. Preferably between 100 and 400.

In a preferred embodiment determination of the conformation score and the binding score are repeated for each pocket and each conformation of the peptide residue in said pocket. The conformation of the peptide is altered by rotating a side chain of the peptide residue by a pre-determined amount. In this way all possible conformations of the peptide side-chain in the pocket can be studied and the best or most likely conformation can be chosen to obtain the binding score.

The conformation of the backbone of the peptide fragment is changed by modelling the conformation of the backbone on any

one of 167 backbones which have been previously generated, based on human and murine crystallographic structures of MHC class II peptide complexes. The backbone conformation and the conformation of the peptide fragment side chains are altered systematically until the conformation score and the binding score of every possible conformation has been determined.

Conveniently the steps are repeated using different peptides from a protein.

In preferred embodiments the binding scores (Y) for different peptides are tabulated and compared. Peptides with the highest scores are predicted to have the highest binding affinity for the particular MHC allele.

In a preferred embodiment the method of determining the binding affinity of a peptide residue for an MHC class II molecule is used in the manufacture of a vaccine derived from a peptide identified by said method.

Preferably the method of determining the binding affinity of a peptide residue for an MHC class II molecule is used to remove potentially immunogenic sequences from a protein and thus reduce said proteins immunogenicity when administered to an organism.

Using the afore-detailed method it is possible to predict the peptides from an auto-immune protein which are presented by MHC class II molecules. Thereafter, it is possible to synthesise peptides which would be antagonists to the presentation of such peptides by the MHC class II molecules. It is also possible to determine any proteins in a vaccine containing heterologous proteins which might result in the stimulation of T cells due to their presentation on MHC class II molecules. These proteins could then be altered or removed depending on their function in the vaccine.

According to a second aspect of the invention there is provided a computer conditioned to receive information characterising a peptide bound to the MHC molecule and to utilise said information to perform a procedure having the following steps;

1) ascertaining the characteristics of a MHC molecule binding groove;

2) presenting a selected peptide, which is selected by a predetermined program, to the MHC molecule and ascertaining a first conformation score;

3) amending the conformation of the peptide, by way of a predetermined program, and ascertaining a second conformation score;

4) repeating step 3 with other conformations of the peptide;

5) selecting the peptide conformation with the highest conformation score; and

6) calculating the binding score from the conformation score.

Preferably the above detailed procedure also includes a step (7) which comprises repeating steps 1-4 with other peptide fragments in the protein to generate information on all peptide fragments in a protein so that a comparison can be made of the strength of the binding between the peptide and the MHC molecule.

Conveniently the above detailed procedure further comprising a step (8) which comprises altering the conformation of the backbone of the peptide fragment.

The use of a computer in such a task is important because there are hundreds of calculations to perform per peptide fragment. A computer conditioned to perform the task can systematically change the conformation of the side chains and the backbone of the peptide fragment while calculating the conformation score and the binding score.

According to a third aspect of the invention there is provided

a pharmaceutical composition made by determining the binding affinity of a peptide for a MHC class II molecule.

A pharmaceutical composition is thus engineered to contain a peptide which is presented by an MHC class II molecule and which therefore stimulates the bodies cellular immune system. Alternatively the pharmaceutical composition is engineered so that it does not include peptides which significantly stimulate the immune system.

The invention will now be described, by way of illustration only, with reference to the following examples, tables and figures accompanying the specification.

Figure 1 shows a graphical representation of the binding score distribution of all 554 13-mer Influenza haemagglutinin peptides bound to HLA-DRB1*0101.

Figure 2 shows a graphical representation of the binding score distribution of all 554 13-mer Influenza haemagglutinin peptides bound to HLA-DRB1*0401.

Table 1 shows the value for all the factors required to determine the binding score for the 15 peptides from Influenza haemagglutinin which have the highest binding affinity for HLA-DRB1*0101.

Table 2 shows the value for all the factors required to determine the binding score for the 15 peptides from Influenza haemagglutinin which have the highest binding affinity for HLA-DRB1*0401.

Table 3 lists the sequence difference between HLA-DRB1*0101 and HLA-DRB1*0401.

Table 4 shows the torsion angles of the mutated side chains in HLA-DRB1*0401.

Example 1

The following method was used to confirm that the peptide PKYVKQNTLKLAT, has a high affinity binding for the MHC molecule HLA-DRB1*0101.

The conformation score was calculated as follows for an oligomeric peptide having thirteen amino acid residues, herein known as a 13-mer peptide:

a) Calculate the steric overlap between the pocket bound peptide residue in the binding groove and an atom forming the pocket; this is value B.

b) Count the number of hydrogen bonds which could be formed between the pocket bound peptide residue and atoms forming the pocket; this is value C.

c) Calculate the strength of electrostatic interactions between any polar atoms of the pocket bound peptide residue and any polar atoms forming the pocket; this is value D.

d) Count the number of favourable contacts between the pocket bound peptide residue and atoms forming the pocket; this is value E.

These values were then transformed into a conformation score (Z) by using the following equation:

$$Z_n=(cK_2C)-(cK_3D)+(cK_4E)-(cK_1B)$$

where cK_1 to cK_4 are constants and n is the number of the pocket. CK_1, cK_2, cK_3 and cK_4 are equal to 100, 1500, 500 and 1000 respectively.

The conformation of each rotatable side chain of the pocket bound peptide bound residue was then altered by 30° and the conformation score was recalculated.

The above steps were repeated for each of the pockets and the

highest conformation score for each of the pockets was used to determine the binding score.

The binding score was determined by establishing values for the following parameters:

a) the number of groove-bound hydrophobic residues; this is value F.

b) the number of non groove-bound hydrophilic residues; this is value G.

c) the number of peptide residues deemed to fit within their respective binding groove; this is value H.

The conformational scores for pockets one and five were doubled and then all the conformational scores were summed to give a value J.

The above values were then imported in to the following equation in order to determine the binding score:

$$J*F^2*(G*H+1)+P$$

The binding scores for all the 13-mer peptides from Influenza Haemagglutinin binding with MHC molecule HLA-DRB1*0401 were calculated and the resultant top 15 binding scores are presented in Table 1. PKYVKQNTLKLAT has the 8th highest binding affinity for HLA-DRB1*0101 from all 554 possible overlapping 13-mer peptides.

Table 1

Rank	Seq.	Peptide	Binding Score	P	B	C	D	E	F	G	H
1	328	NTLKLATGMRNVP	9382500	15012	0.00	1		27	4	6	5
2	453	IDLTDSEMNKLFE	8288922	17964	0.72	1		40	3	6	5
3	373	NSEGTGQAADLKS	7520420	10661	0.68	0	+0.01	30	4	7	
4	504	HDVYRDEALNNRF	7211042	15527	0.56	1	-0.05	31	3	6	5
5	119	PDYASLRSLVASS	7174962	17351	0.68	1		40	4	4	5
6	461	NKLFEKTRRQLRE	7049469	19407	0.79	0	+0.01	56	2	7	5
7	122	ASLRSLVASSGTL	6922064	16346	0.09	0		25	4	4	5
8	322	PKYVKQNTLKLAT	6765975	18217	1.82	1		56	3	5	5
9	458	SEMNKLFEKTRRQ	6156822	16617	0.30	4	+0.08	44	2	7	5
10	513	NNRFQIKGVELKS	6096900	14052	1.32	3	-0.01	30	4	7	4
11	439	YNAELLVALENQH	5890199	14198	0.60	1		33	4	4	5
12	63	STGKICNNPHRIL	5887908	12776	0.75	5	-0.05	31	3	6	5
13	50	IEVTNATELVQSS	5503551	14297	0.95	2	+0.06	39	3	5	5
14	262	NSNGNLIAPRGYF	5284475	10102	0.09	1		21	4	5	5
15	257	DVLVINSNGNLIA	5239292	17028	1.35	2		35	3	4	5

Example 2

A method as described in Example 1 was used to confirm that the peptide PDYASLRSLVASS from Influenza haemagglutinin, has high affinity binding for the MHC molecule HLA-DRB1*0401.

The structure of HLA-DRB1*0401 is not known but a three dimensional model was constructed based on the known structure of HLA-DRB1*0101 by homology modelling. 10 amino acid differences between the two molecules were identified (see Table 2) and HLA-DRB1*0101 was mutated using the molecular modelling package 'Quanta' to produce a model of HLA-DRB1*0401.

Then the side-chain conformations of the 10 amino acids were adjusted interactively. In most cases, torsion angles were chosen which resulted in little or no steric overlap between the mutated residues and surrounding atoms. In the case of non-conserved residues which were either charged or whose side-chains were able to form hydrogen bonds, the potential to form favourable interactions was also considered. The placement of 13H, 28D and 71K was such that these residues were able to form a favourable electrostatic arrangement whilst at the same time, having minimum steric overlap with surrounding atoms. In the case of 30Y, this residue was positioned such that its hydroxyl group was situated close to the side-chain of 9E, where a hydrogen bond may be formed. The torsion angles chosen for the 10 mutated amino acid residues were calculated in accordance with the standard conventions and are listed in Table 3.

The binding scores for all 13-mer peptides from Influenza Haemagglutinin binding with MHC molecule HLA-DRB1*0401 were calculated and the resultant top 15 binding scores are presented in Table 4. PDYASLRSLVASS has the 9th highest binding affinity for HLA-DRB1*0401 from all 554 possible overlapping 13-mer peptides.

Table 2

Seq. Pos.	HLA-DRB1*0101	HLA-DRB1*0401
b9	Tryptophan	Glutamic acid
b11	Leucine	Valine
b13	Phenylalanine	Histidine
b26	Leucine	Phenylalanine
b28	Glutamic acid	Aspartic Acid
b30	Cysteine	Tyrosine
b31	Isoleucine	Phenylalanine
b33	Asparagine	Histidine
b37	Serine	Tyrosine
b71	Arginine	Lysine

Table 3

Residue	c1	c2	c3	c4
b9	-61°	-71°	-2°	
b11	168°			
b13	-38°	-63°		
b26	170°	57°		
b28	-174°	-15°		
b30	-174°	41°		
b31	-119°	-13°		
b33	-95°	-2°		
b37	-116°	-2°		
b71	-97°	-45°	172°	9°

Table 4

Rank	Seq.	Peptide	Binding Score	P	B	C	D	E	F	G	H
1	453	IDLTDSEMNKLFE	3070823	6559	0.36	0		42	3	6	5
2	373	NSEGTGQAADLKS	2988447	4182	0.36	0	+0.01	32	4	7	5
3	328	NTLKLATGMRNVP	2899375	4639	0.00	1		27	4	6	5
4	122	ASLRSLVASSGTL	2894599	6819	0.03	0		24	4	4	5
5	72	HRILDGIDCTLID	2820446	4623	0.60	1	+0.16	28	4	6	5
6	461	NKLFEKTRRQLRE	2662369	7203	0.36	0	-0.11	50	2	7	5
7	119	PDYASLRSLVASS	2616648	6184	0.11	1		32	4	4	5
8	188	DNFDKLYIWGIHH	2615259	5429	0.58	0		29	5	6	4
9	322	PKYVKQNTLKLAT	2515861	6407	0.46	2		44	3	5	5
10	232	NIGSRPWVRGLSS	2488137	4818	0.41	0	-0.02	35	4	5	5
11	504	HDVYRDEALNNRF	2353661	4965	0.05	1	-0.07	25	3	6	5
12	135	EFITEGFTWTGVT	2208179	3543	0.07	1		20	4	5	5
13	251	TIVKPGDVLVINS	2176819	5259	0.10	0		16	5	5	4
14	257	DVLVINSNGNLIA	2107570	6673	0.71	2		40	3	4	5
15	439	YNAELLVALENQH	2035430	4795	0.03	1		26	4	4	5

Example 3

A library of backbones were constructed by examining the crystal structure of the HLA-DR1 complexed with SEB super-antigen. This results in a collection of homogenous peptides within the MHC binding groove. The atomic positions of the peptide backbone, as shown in the PDB file produced from the crystal, were considered to be the `representative' backbone conformation of a peptide which binds to HLA-DR1.

Each of the peptide backbone conformations from the known MHC class II crystallographic structures are taken and after being transformed to the same frame of reference as the `representative' peptide had the differences between their Cα/Cβ positions and those of the `representative' peptide

calculated. These differences summarise the variability of Cα/Cβ atomic positions between the known peptides and the `representative' peptide.

The differences were doubled to take into account the fact that the variability of peptides thus far crystallised may not fully represent the true variability of peptides binding to MHC class II molecules. The differences were then used to define regions within which peptide Cα and Cβ atoms centres are constrained to lie.

An exhaustive search was then made through candidate peptide backbones. Starting from the `representative' peptide candidates are generated by adjusting backbone ϕ and ψ angles in ten degree steps from the N-terminus to the C-terminus. An adjustment was rejected if it led to any Cα or Cβ atom centre being outside the allowed region, derived above. An adjustment which did not violate the constraint results in a new backbone conformation which is stored within the peptide backbone library.

The x, y, and z co-ordinates of atoms in the backbones designated 0, 14, 62, 65, 75, 93, 104, 107, 112, 118, 129, 134, 141, 144 are given in Tables 5 to 18.

Table 5

Backbone 0					
Atom Number	Atom type	Position in peptide	x	y	z
0	N	0	19.913	86.191	20.687
1	CA	0	19.472	86.222	22.078
2	C	0	18.153	85.531	22.516
3	O	0	18.200	84.640	23.352
4	CB	0	19.504	87.660	22.593
5	N	1	16.984	85.957	22.044
6	CA	1	15.771	85.316	22.536
7	C	1	15.262	84.115	21.770
8	O	1	15.175	84.127	20.547
9	CB	1	14.663	86.325	22.743
10	N	2	14.959	83.055	22.510
11	CA	2	14.414	81.829	21.926
12	C	2	12.920	82.131	21.907
13	O	2	12.384	82.737	22.840
14	CB	2	14.756	80.548	22.811
15	N	3	12.283	81.841	20.784
16	CA	3	10.866	82.097	20.637
17	C	3	10.086	80.785	20.839
18	O	3	10.560	79.730	20.447
19	CB	3	10.624	82.744	19.230
20	N	4	8.951	80.855	21.528
21	CA	4	8.035	79.734	21.814
22	C	4	6.945	79.658	20.721
23	O	4	6.664	80.648	20.044
24	CB	4	7.330	79.991	23.185
25	N	5	6.355	78.499	20.461
26	CA	5	5.266	78.527	19.496
27	C	5	4.167	78.292	20.475
28	O	5	4.342	77.560	21.444
29	CB	5	5.349	77.437	18.471
30	N	6	3.044	78.938	20.261
31	CA	6	1.950	78.858	21.205
32	C	6	1.050	77.758	20.856
33	O	6	0.836	77.517	19.690
34	CB	6	1.163	80.226	21.247
35	N	7	0.420	77.190	21.863
36	CA	7	-0.503	76.102	21.660
37	C	7	-1.889	76.607	21.227
38	O	7	-2.429	77.551	21.833
39	CB	7	-0.611	75.340	22.937
40	N	8	-2.442	75.997	20.167
41	CA	8	-3.790	76.330	19.644

Table 5 continued

Atom Number	Atom type	Position in peptide	x	y	z
42	C	8	-4.839	75.618	20.504
43	O	8	-4.505	74.687	21.236
44	CB	8	-3.924	75.908	18.149
45	N	9	-6.093	76.041	20.436
46	CA	9	-7.113	75.382	21.236
47	C	9	-7.976	74.424	20.403
48	O	9	-8.366	74.742	19.266
49	CB	9	-7.963	76.413	21.973
50	N	10	-8.203	73.232	20.971
51	CA	10	-8.995	72.149	20.365
52	C	10	-10.492	72.527	20.200
53	O	10	-10.962	73.563	20.702
54	CB	10	-8.830	70.835	21.191
55	N	11	-11.238	71.661	19.523
56	CA	11	-12.654	71.907	19.270
57	C	11	-13.603	71.483	20.395
58	O	11	-13.661	70.302	20.800
59	CB	11	-13.072	71.269	17.940
60	N	12	-14.360	72.481	20.852
61	CA	12	-15.363	72.337	21.898
62	C	12	-14.758	72.166	23.281
63	O	12	-14.785	71.069	23.853
64	CB	12	-16.320	71.168	21.577

Table 6

Backbone 14					
Atom Number	Atom type	Position in peptide	x	y	z
0	N	0	0.000	0.000	0.000
1	CA	0	18.281	86.637	22.405
2	C	0	16.799	86.756	22.715
3	O	0	16.250	87.880	22.720
4	CB	0	0.000	0.000	0.000
5	N	1	16.174	85.601	22.931
6	CA	1	14.768	85.553	23.287
7	C	1	14.098	84.393	22.569
8	O	1	13.053	84.588	21.908
9	CB	1	14.090	86.846	22.869
10	N	2	14.723	83.223	22.680
11	CA	2	14.182	82.013	22.093
12	C	2	12.659	82.164	21.901
13	O	2	11.952	82.431	22.884
14	CB	2	14.470	80.825	22.994
15	N	3	12.242	82.022	20.649
16	CA	3	10.845	82.086	20.317
17	C	3	10.219	80.681	20.423
18	O	3	10.898	79.694	20.101
19	CB	3	10.669	82.621	18.906
20	N	4	8.980	80.660	20.898
21	CA	4	8.245	79.430	21.010
22	C	4	6.863	79.586	20.344
23	O	4	6.283	80.680	20.413
24	CB	4	8.071	79.059	22.472
25	N	5	6.427	78.504	19.710
26	CA	5	5.135	78.479	19.082
27	C	5	4.084	77.942	20.074
28	O	5	4.171	76.770	20.468
29	CB	5	5.174	77.593	17.848
30	N	6	3.174	78.832	20.452
31	CA	6	2.100	78.470	21.336
32	C	6	1.349	77.248	20.769
33	O	6	1.703	76.776	19.678
34	CB	6	1.139	79.635	21.492
35	N	7	0.381	76.781	21.550
36	CA	7	-0.441	75.677	21.137
37	C	7	-1.906	76.139	21.008
38	O	7	-2.505	76.533	22.020
39	CB	7	-0.346	74.551	22.153
40	N	8	-2.392	76.101	19.773
41	CA	8	-3.758	76.454	19.498
42	C	8	-4.704	75.537	20.299
43	O	8	-4.316	74.404	20.618
44	CB	8	-4.043	76.313	18.013

Table 6 continued

Atom Number	Atom type	Position in peptide	x	y	z
45	N	9	-5.873	76.084	20.610
46	CA	9	-6.881	75.338	21.313
47	C	9	-7.500	74.285	20.371
48	O	9	-7.243	74.336	19.159
49	CB	9	-7.964	76.275	21.818
50	N	10	-8.250	73.372	20.978
51	CA	10	-8.934	72.354	20.229
52	C	10	-10.393	72.786	19.976
53	O	10	-11.075	73.192	20.928
54	CB	10	-8.914	71.043	20.996
55	N	11	-10.781	72.710	18.708
56	CA	11	-12.127	73.032	18.320
57	C	11	-13.058	71.846	18.640
58	O	11	-13.254	70.984	17.770
59	CB	11	-12.180	73.341	16.834
60	N	12	-13.551	71.844	19.872
61	CA	12	-14.474	70.830	20.305
62	C	12	0.000	-12.127	73.032
63	O	12	18.356	0.000	-12.127
64	CB	12	0.000	0.000	0.000

Table 7

Backbone 62					
Atom Number	Atom type	Position in peptide	x	y	z
0	N	0	0.000	0.000	0.000
1	CA	0	18.315	86.971	22.396
2	C	0	16.796	86.979	22.404
3	O	0	16.173	87.867	21.780
4	CB	0	0.000	0.000	0.000
5	N	1	16.231	85.979	23.075
6	CA	1	14.791	85.876	23.216
7	C	1	14.286	84.665	22.451
8	O	1	13.659	84.820	21.380
9	CB	1	14.132	87.123	22.652
10	N	2	14.595	83.487	22.989
11	CA	2	14.144	82.241	22.404
12	C	2	12.614	82.280	22.212
13	O	2	11.890	82.495	23.195
14	CB	2	14.518	81.077	23.305
15	N	3	12.208	82.108	20.960
16	CA	3	10.810	82.071	20.629
17	C	3	10.289	80.623	20.734
18	O	3	11.105	79.691	20.783
19	CB	3	10.596	82.591	19.218
20	N	4	8.967	80.514	20.800
21	CA	4	8.328	79.228	20.852
22	C	4	6.861	79.356	20.395
23	O	4	6.157	80.256	20.876
24	CB	4	8.377	78.680	22.268
25	N	5	6.490	78.478	19.470
26	CA	5	5.140	78.440	18.978
27	C	5	4.171	78.141	20.139
28	O	5	4.543	77.392	21.055
29	CB	5	5.006	77.369	17.909
30	N	6	3.002	78.765	20.060
31	CA	6	1.975	78.549	21.042
32	C	6	1.039	77.416	20.577
33	O	6	1.276	76.842	19.503
34	CB	6	1.174	79.824	21.246
35	N	7	0.052	77.131	21.418
36	CA	7	-0.931	76.132	21.102
37	C	7	-2.325	76.784	21.008
38	O	7	-2.553	77.814	21.661
39	CB	7	-0.941	75.055	22.174
40	N	8	-3.166	76.177	20.179
41	CA	8	-4.518	76.638	20.020
42	C	8	-5.491	75.631	20.666
43	O	8	-5.155	74.441	20.754

Table 7 continued

Atom Number	Atom type	Position in peptide	x	y	z
44	CB	8	-4.845	76.793	18.545
45	N	9	-6.623	76.163	21.113
46	CA	9	-7.650	75.345	21.696
47	C	9	-8.161	74.329	20.655
48	O	9	-8.197	74.658	19.460
49	CB	9	-8.802	76.215	22.170
50	N	10	-8.492	73.143	21.153
51	CA	10	-9.030	72.107	20.315
52	C	10	-10.518	72.390	20.029
53	O	10	-11.258	72.730	20.964
54	CB	10	-8.887	70.758	21.000
55	N	11	-10.869	72.271	18.754
56	CA	11	-12.232	72.455	18.336
57	C	11	-13.047	71.182	18.641
58	O	11	-13.155	70.312	17.764
59	CB	11	-12.284	72.752	16.847
60	N	12	-13.544	71.124	19.871
61	CA	12	-14.366	70.022	20.291
62	C	12	0.000	-12.232	72.455
63	O	12	18.332	0.000	-12.232
64	CB	12	0.000	0.000	0.000

Table 8

Backbone 65					
Atom Number	Atom type	Position in peptide	x	y	z
0	N	0	0.000	0.000	0.000
1	CA	0	18.487	86.641	22.418
2	C	0	16.990	86.870	22.533
3	O	0	16.510	87.999	22.287
4	CB	0	0.000	0.000	0.000
5	N	1	16.279	85.796	22.868
6	CA	1	14.844	85.866	23.065
7	C	1	14.178	84.664	22.417
8	O	1	13.234	84.830	21.612
9	CB	1	14.301	87.132	22.424
10	N	2	14.699	83.484	22.746
11	CA	2	14.144	82.241	22.248
12	C	2	12.616	82.381	22.089
13	O	2	11.950	82.822	23.038
14	CB	2	14.457	81.109	23.212
15	N	3	12.150	82.035	20.895
16	CA	3	10.742	82.065	20.608
17	C	3	10.206	80.624	20.484
18	O	3	10.895	79.773	19.902
19	CB	3	10.491	82.818	19.314
20	N	4	9.029	80.419	21.065
21	CA	4	8.376	79.140	20.993
22	C	4	6.930	79.322	20.491
23	O	4	6.309	80.350	20.801
24	CB	4	8.365	78.486	22.364
25	N	5	6.484	78.339	19.718
26	CA	5	5.139	78.340	19.212
27	C	5	4.150	78.069	20.363
28	O	5	4.487	77.306	21.280
29	CB	5	4.985	77.274	18.142
30	N	6	3.002	78.731	20.275
31	CA	6	1.959	78.547	21.246
32	C	6	0.861	77.634	20.665
33	O	6	0.752	77.533	19.433
34	CB	6	1.360	79.890	21.628
35	N	7	0.134	76.994	21.573
36	CA	7	-0.959	76.143	21.187
37	C	7	-1.983	76.952	20.366
38	O	7	-1.708	78.116	20.039
39	CB	7	-1.631	75.569	22.422
40	N	8	-3.087	76.287	20.048
41	CA	8	-4.156	76.921	19.326
42	C	8	-5.496	76.242	19.676

Table 8 continued

Atom Number	Atom type	Position in peptide	x	y	z
43	O	8	-6.146	75.692	18.775
44	CB	8	-3.906	76.820	17.831
45	N	9	-5.817	76.283	20.964
46	CA	9	-7.058	75.736	21.439
47	C	9	-7.606	74.721	20.416
48	O	9	-7.311	74.855	19.219
49	CB	9	-8.071	76.849	21.649
50	N	10	-8.339	73.746	20.940
51	CA	10	-8.959	72.751	20.108
52	C	10	-10.421	73.147	19.824
53	O	10	-10.685	73.773	18.787
54	CB	10	-8.919	71.398	20.799
55	N	11	-11.294	72.734	20.735
56	CA	11	-12.689	73.067	20.635
57	C	11	-13.474	71.860	20.085
58	O	11	-13.031	71.253	19.099
59	CB	11	-12.873	74.262	19.715
60	N	12	-14.572	71.556	20.766
61	CA	12	-15.436	70.486	20.348
62	C	12	0.000	-12.689	73.067
63	O	12	18.675	0.000	-12.689
64	CB	12	0.000	0.000	0.000

Table 9

Backbone 75					
Atom Number	Atom type	Position in peptide	x	y	z
0	N	0	0.000	0.000	0.000
1	CA	0	18.442	86.539	22.377
2	C	0	16.947	86.419	22.136
3	O	0	16.452	86.839	21.066
4	CB	0	0.000	0.000	0.000
5	N	1	16.265	85.822	23.109
6	CA	1	14.823	85.676	23.048
7	C	1	14.466	84.417	22.277
8	O	1	14.197	84.487	21.057
9	CB	1	14.218	86.875	22.338
10	N	2	14.505	83.290	22.985
11	CA	2	14.144	82.013	22.404
12	C	2	12.615	81.942	22.214
13	O	2	11.895	81.727	23.200
14	CB	2	14.601	80.882	23.308
15	N	3	12.201	82.159	20.971
16	CA	3	10.808	82.078	20.626
17	C	3	10.331	80.615	20.726
18	O	3	11.176	79.709	20.772
19	CB	3	10.592	82.592	19.213
20	N	4	9.013	80.465	20.789
21	CA	4	8.414	79.160	20.836
22	C	4	6.944	79.245	20.377
23	O	4	6.322	80.304	20.544
24	CB	4	8.478	78.609	22.251
25	N	5	6.482	78.145	19.793
26	CA	5	5.116	78.053	19.354
27	C	5	4.181	77.969	20.577
28	O	5	4.609	77.470	21.629
29	CB	5	4.932	76.823	18.483
30	N	6	2.974	78.490	20.389
31	CA	6	1.974	78.445	21.420
32	C	6	0.736	77.679	20.910
33	O	6	0.349	77.867	19.748
34	CB	6	1.576	79.855	21.821
35	N	7	0.206	76.836	21.788
36	CA	7	-0.980	76.086	21.478
37	C	7	-1.844	76.872	20.470
38	O	7	-1.448	77.977	20.071
39	CB	7	-1.778	75.828	22.745
40	N	8	-2.952	76.249	20.088
41	CA	8	-3.885	76.873	19.189

Table 9 continued

Atom Number	Atom type	Position in peptide	x	y	z
42	C	8	-5.324	76.483	19.579
43	O	8	-6.195	76.435	18.698
44	CB	8	-3.604	76.435	17.762
45	N	9	-5.491	76.194	20.865
46	CA	9	-6.786	75.859	21.391
47	C	9	-7.424	74.747	20.535
48	O	9	-7.209	74.729	19.314
49	CB	9	-7.681	77.087	21.388
50	N	10	-8.142	73.864	21.219
51	CA	10	-8.840	72.797	20.556
52	C	10	-10.312	73.196	20.334
53	O	10	-10.616	73.833	19.314
54	CB	10	-8.772	71.532	21.394
55	N	11	-11.149	72.774	21.275
56	CA	11	-12.546	73.108	21.233
57	C	11	-13.321	72.011	20.475
58	O	11	-12.815	71.509	19.460
59	CB	11	-12.741	74.445	20.540
60	N	12	-14.483	71.674	21.023
61	CA	12	-15.343	70.702	20.406
62	C	12	0.000	-12.546	73.108
63	O	12	18.817	0.000	-12.546
64	CB	12	0.000	0.000	0.000

Table 10

Backbone 93					
Atom Number	Atom type	Position in peptide	x	y	z
0	N	0	0.000	0.000	0.000
1	CA	0	18.249	86.312	21.629
2	C	0	16.910	86.341	22.345
3	O	0	16.646	87.271	23.139
4	CB	0	0.000	0.000	0.000
5	N	1	16.080	85.351	22.027
6	CA	1	14.782	85.213	22.662
7	C	1	14.078	83.978	22.127
8	O	1	12.999	84.095	21.505
9	CB	1	13.932	86.434	22.357
10	N	2	14.712	82.828	22.345
11	CA	2	14.144	81.558	21.938
12	C	2	12.613	81.689	21.812
13	O	2	11.912	81.568	22.828
14	CB	2	14.484	80.486	22.959
15	N	3	12.179	81.964	20.587
16	CA	3	10.775	82.068	20.300
17	C	3	10.163	80.658	20.176
18	O	3	10.712	79.826	19.439
19	CB	3	10.564	82.834	19.005
20	N	4	9.085	80.454	20.925
21	CA	4	8.374	79.206	20.882
22	C	4	7.026	79.401	20.159
23	O	4	6.568	80.546	20.036
24	CB	4	8.130	78.697	22.292
25	N	5	6.482	78.283	19.690
26	CA	5	5.203	78.295	19.035
27	C	5	4.087	78.033	20.066
28	O	5	4.298	77.235	20.991
29	CB	5	5.163	77.229	17.954
30	N	6	2.980	78.741	19.876
31	CA	6	1.833	78.572	20.726
32	C	6	1.164	77.213	20.434
33	O	6	1.603	76.513	19.510
34	CB	6	0.839	79.695	20.486
35	N	7	0.169	76.899	21.254
36	CA	7	-0.585	75.687	21.080
37	C	7	-2.092	76.013	21.037
38	O	7	-2.667	76.338	22.086
39	CB	7	-0.300	74.729	22.223
40	N	8	-2.639	75.944	19.829
41	CA	8	-4.045	76.173	19.635
42	C	8	-4.853	75.344	20.653
43	O	8	-4.314	74.368	21.198

Table 10 continued

Atom Number	Atom type	Position in peptide	x	y	z
44	CB	8	-4.445	75.782	18.223
45	N	9	-6.082	75.791	20.882
46	CA	9	-6.974	75.097	21.769
47	C	9	-8.018	74.312	20.948
48	O	9	-8.754	74.928	20.163
49	CB	9	-7.679	76.089	22.679
50	N	10	-8.002	72.999	21.144
51	CA	10	-8.947	72.137	20.488
52	C	10	-10.274	72.891	20.269
53	O	10	-10.348	73.727	19.356
54	CB	10	-9.194	70.899	21.332
55	N	11	-11.256	72.533	21.087
56	CA	11	-12.539	73.179	21.038
57	C	11	-13.542	72.288	20.278
58	O	11	-13.224	71.836	19.167
59	CB	11	-12.418	74.524	20.343
60	N	12	-14.678	72.054	20.925
61	CA	12	-15.731	71.281	20.326
62	C	12	0.000	-12.539	73.179
63	O	12	18.616	0.000	-12.539
64	CB	12	0.000	0.000	0.000

Table 11

Backbone 104					
Atom Number	Atom type	Position in peptide	x	y	z
0	N	0	0.000	0.000	0.000
1	CA	0	18.400	86.585	22.355
2	C	0	16.914	86.850	22.523
3	O	0	16.453	87.991	22.296
4	CB	0	0.000	0.000	0.000
5	N	1	16.189	85.793	22.880
6	CA	1	14.763	85.897	23.128
7	C	1	14.059	84.662	22.593
8	O	1	12.980	84.778	21.971
9	CB	1	14.210	87.122	22.421
10	N	2	14.693	83.511	22.810
11	CA	2	14.125	82.241	22.404
12	C	2	12.594	82.372	22.277
13	O	2	11.945	82.807	23.241
14	CB	2	14.465	81.169	23.424
15	N	3	12.104	82.026	21.093
16	CA	3	10.690	82.048	20.837
17	C	3	10.159	80.604	20.723
18	O	3	10.919	79.713	20.317
19	CB	3	10.406	82.801	19.548
20	N	4	8.902	80.444	21.120
21	CA	4	8.250	79.166	21.029
22	C	4	6.905	79.319	20.290
23	O	4	6.415	80.450	20.160
24	CB	4	8.009	78.605	22.420
25	N	5	6.401	78.185	19.817
26	CA	5	5.130	78.158	19.147
27	C	5	4.011	77.862	20.165
28	O	5	4.164	76.935	20.975
29	CB	5	5.135	77.091	18.066
30	N	6	2.968	78.680	20.096
31	CA	6	1.823	78.502	20.947
32	C	6	1.166	77.138	20.656
33	O	6	1.718	76.360	19.864
34	CB	6	0.819	79.617	20.708
35	N	7	0.047	76.906	21.334
36	CA	7	-0.707	75.699	21.135
37	C	7	-2.213	76.030	21.083
38	O	7	-2.793	76.357	22.129
39	CB	7	-0.435	74.724	22.267
40	N	8	-2.754	75.961	19.873
41	CA	8	-4.157	76.194	19.670
42	C	8	-4.974	75.368	20.684
43	O	8	-4.444	74.387	21.228

Table 11 continued

Atom Number	Atom type	Position in peptide	x	y	z
44	CB	8	-4.550	75.803	18.256
45	N	9	-6.200	75.824	20.911
46	CA	9	-7.100	75.134	21.794
47	C	9	-8.146	74.358	20.969
48	O	9	-8.997	74.991	20.328
49	CB	9	-7.800	76.129	22.704
50	N	10	-8.007	73.038	21.000
51	CA	10	-8.934	72.175	20.320
52	C	10	-10.266	72.919	20.092
53	O	10	-10.341	73.752	19.177
54	CB	10	-9.181	70.924	21.145
55	N	11	-11.249	72.557	20.907
56	CA	11	-12.537	73.194	20.850
57	C	11	-13.529	72.294	20.086
58	O	11	-13.514	72.297	18.847
59	CB	11	-12.421	74.537	20.152
60	N	12	-14.310	71.549	20.860
61	CA	12	-15.320	70.695	20.297
62	C	12	0.000	-12.537	73.194
63	O	12	18.422	0.000	-12.537
64	CB	12	0.000	0.000	0.000

Table 12

Backbone 107					
Atom Number	Atom type	Position in peptide	x	y	z
0	N	0	0.000	0.000	0.000
1	CA	0	18.468	86.641	22.418
2	C	0	16.971	86.870	22.533
3	O	0	16.491	87.999	22.287
4	CB	0	0.000	0.000	0.000
5	N	1	16.260	85.796	22.868
6	CA	1	14.825	85.866	23.065
7	C	1	14.159	84.664	22.417
8	O	1	13.215	84.830	21.612
9	CB	1	14.282	87.132	22.424
10	N	2	14.680	83.484	22.746
11	CA	2	14.125	82.241	22.248
12	C	2	12.597	82.381	22.089
13	O	2	11.931	82.822	23.038
14	CB	2	14.438	81.109	23.212
15	N	3	12.131	82.035	20.895
16	CA	3	10.723	82.065	20.608
17	C	3	10.187	80.624	20.484
18	O	3	10.876	79.773	19.902
19	CB	3	10.472	82.818	19.314
20	N	4	9.010	80.419	21.065
21	CA	4	8.357	79.140	20.993
22	C	4	6.911	79.322	20.491
23	O	4	6.290	80.350	20.801
24	CB	4	8.346	78.486	22.364
25	N	5	6.465	78.339	19.718
26	CA	5	5.120	78.340	19.212
27	C	5	4.131	78.069	20.363
28	O	5	4.469	77.306	21.280
29	CB	5	4.966	77.274	18.142
30	N	6	2.983	78.731	20.275
31	CA	6	1.940	78.547	21.246
32	C	6	0.842	77.634	20.665
33	O	6	0.733	77.533	19.433
34	CB	6	1.341	79.890	21.628
35	N	7	0.115	76.994	21.573
36	CA	7	-0.978	76.143	21.187
37	C	7	-2.002	76.952	20.366
38	O	7	-1.726	78.116	20.039
39	CB	7	-1.650	75.569	22.422
40	N	8	-3.106	76.287	20.048
41	CA	8	-4.175	76.921	19.326
42	C	8	-5.514	76.242	19.676
43	O	8	-6.165	75.692	18.775

Table 12 continued

Atom Number	Atom type	Position in peptide	x	y	z
44	CB	8	-3.925	76.820	17.831
45	N	9	-5.836	76.283	20.964
46	CA	9	-7.077	75.736	21.439
47	C	9	-7.625	74.721	20.416
48	O	9	-7.330	74.855	19.219
49	CB	9	-8.090	76.849	21.649
50	N	10	-8.358	73.746	20.940
51	CA	10	-8.977	72.751	20.108
52	C	10	-10.440	73.147	19.824
53	O	10	-10.703	73.773	18.787
54	CB	10	-8.938	71.398	20.799
55	N	11	-11.313	72.734	20.735
56	CA	11	-12.708	73.067	20.635
57	C	11	-13.493	71.860	20.085
58	O	11	-13.050	71.253	19.099
59	CB	11	-12.892	74.262	19.715
60	N	12	-14.591	71.556	20.766
61	CA	12	-15.455	70.486	20.348
62	C	12	0.000	-12.708	73.067
63	O	12	18.675	0.000	-12.708
64	CB	12	0.000	0.000	0.000

Table 13

Backbone 112					
Atom Number	Atom type	Position in peptide	x	y	z
0	N	0	0.000	0.000	0.000
1	CA	0	18.408	86.726	22.399
2	C	0	16.919	86.606	22.121
3	O	0	16.449	87.028	21.041
4	CB	0	0.000	0.000	0.000
5	N	1	16.215	86.005	23.077
6	CA	1	14.774	85.858	22.981
7	C	1	14.438	84.649	22.125
8	O	1	14.190	84.795	20.907
9	CB	1	14.176	87.097	22.337
10	N	2	14.470	83.480	22.761
11	CA	2	14.125	82.241	22.093
12	C	2	12.600	82.176	21.872
13	O	2	11.849	82.152	22.858
14	CB	2	14.572	81.057	22.932
15	N	3	12.224	82.187	20.598
16	CA	3	10.839	82.083	20.230
17	C	3	10.319	80.669	20.557
18	O	3	11.133	79.744	20.692
19	CB	3	10.674	82.359	18.745
20	N	4	9.001	80.583	20.701
21	CA	4	8.361	79.323	20.960
22	C	4	6.868	79.411	20.585
23	O	4	6.126	80.158	21.239
24	CB	4	8.500	78.961	22.429
25	N	5	6.516	78.676	19.537
26	CA	5	5.150	78.615	19.095
27	C	5	4.229	78.301	20.291
28	O	5	4.706	77.734	21.285
29	CB	5	4.995	77.540	18.033
30	N	6	2.976	78.716	20.149
31	CA	6	1.986	78.455	21.158
32	C	6	0.948	77.449	20.621
33	O	6	1.060	77.031	19.459
34	CB	6	1.291	79.747	21.552
35	N	7	0.020	77.088	21.499
36	CA	7	-1.045	76.194	21.133
37	C	7	-2.219	76.999	20.540
38	O	7	-2.062	78.205	20.301
39	CB	7	-1.517	75.422	22.353
40	N	8	-3.314	76.286	20.301
41	CA	8	-4.508	76.904	19.793
42	C	8	-5.720	75.987	20.056
43	O	8	-5.881	74.984	19.345
44	CB	8	-4.369	77.156	18.302
45	N	9	-6.483	76.357	21.078

Table 13 continued

Atom Number	Atom type	Position in peptide	x	y	z
46	CA	9	-7.676	75.631	21.417
47	C	9	-7.858	74.446	20.447
48	O	9	-7.297	74.482	19.341
49	CB	9	-8.883	76.549	21.338
50	N	10	-8.598	73.451	20.920
51	CA	10	-8.898	72.298	20.116
52	C	10	-10.415	72.236	19.842
53	O	10	-11.204	72.400	20.784
54	CB	10	-8.455	71.034	20.832
55	N	11	-10.740	72.040	18.569
56	CA	11	-12.112	71.910	18.163
57	C	11	-12.689	70.583	18.695
58	O	11	-12.384	69.523	18.128
59	CB	11	-12.211	71.942	16.648
60	N	12	-13.459	70.705	19.770
61	CA	12	-14.109	69.563	20.354
62	C	12	0.000	-12.112	71.910
63	O	12	18.708	0.000	-12.112
64	CB	12	0.000	0.000	0.000

Table 14

Backbone 118					
Atom Number	Atom type	Position in peptide	x	y	z
0	N	0	0.000	0.000	0.000
1	CA	0	18.471	86.536	22.407
2	C	0	16.968	86.701	22.266
3	O	0	16.498	87.742	21.755
4	CB	0	0.000	0.000	0.000
5	N	1	16.246	85.665	22.686
6	CA	1	14.795	85.690	22.663
7	C	1	14.271	84.435	21.986
8	O	1	13.620	84.525	20.922
9	CB	1	14.318	86.904	21.884
10	N	2	14.591	83.292	22.589
11	CA	2	14.125	82.013	22.093
12	C	2	12.591	82.045	21.934
13	O	2	11.881	82.067	22.951
14	CB	2	14.518	80.907	23.057
15	N	3	12.165	82.081	20.677
16	CA	3	10.762	82.064	20.366
17	C	3	10.221	80.625	20.479
18	O	3	11.005	79.674	20.343
19	CB	3	10.536	82.588	18.958
20	N	4	8.925	80.541	20.756
21	CA	4	8.263	79.268	20.845
22	C	4	6.879	79.352	20.171
23	O	4	6.325	80.457	20.070
24	CB	4	8.101	78.868	22.301
25	N	5	6.413	78.195	19.716
26	CA	5	5.115	78.103	19.106
27	C	5	4.061	77.755	20.177
28	O	5	4.217	76.737	20.866
29	CB	5	5.122	77.034	18.027
30	N	6	3.069	78.632	20.282
31	CA	6	1.984	78.421	21.202
32	C	6	1.060	77.308	20.670
33	O	6	1.327	76.771	19.584
34	CB	6	1.192	79.706	21.374
35	N	7	0.048	76.997	21.472
36	CA	7	-0.928	76.012	21.093
37	C	7	-2.316	76.673	20.976
38	O	7	-2.546	77.708	21.619
39	CB	7	-0.975	74.902	22.128
40	N	8	-3.150	76.066	20.139
41	CA	8	-4.496	76.535	19.959
42	C	8	-5.484	75.538	20.596
43	O	8	-5.163	74.343	20.680
44	CB	8	-4.801	76.684	18.479

Table 14 continued

Atom Number	Atom type	Position in peptide	x	y	z
45	N	9	-6.612	76.081	21.040
46	CA	9	-7.652	75.273	21.615
47	C	9	-8.169	74.268	20.567
48	O	9	-8.200	74.604	19.374
49	CB	9	-8.795	76.156	22.087
50	N	10	-8.513	73.083	21.059
51	CA	10	-9.059	72.056	20.214
52	C	10	-10.544	72.355	19.925
53	O	10	-11.281	72.703	20.859
54	CB	10	-8.931	70.703	20.892
55	N	11	-10.894	72.239	18.649
56	CA	11	-12.254	72.439	18.229
57	C	11	-13.135	71.287	18.754
58	O	11	-13.091	70.187	18.183
59	CB	11	-12.328	72.490	16.713
60	N	12	-13.856	71.586	19.828
61	CA	12	-14.763	70.632	20.406
62	C	12	0.000	-12.254	72.439
63	O	12	18.754	0.000	-12.254
64	CB	12	0.000	0.000	0.000

Table 15

Backbone 129					
Atom Number	Atom type	Position in peptide	x	y	z
0	N	0	0.000	0.000	0.000
1	CA	0	18.495	86.291	22.091
2	C	0	17.099	86.364	22.686
3	O	0	16.668	87.449	23.137
4	CB	0	0.000	0.000	0.000
5	N	1	16.409	85.228	22.645
6	CA	1	15.079	85.125	23.217
7	C	1	14.331	83.972	22.570
8	O	1	13.400	84.204	21.766
9	CB	1	14.313	86.412	22.964
10	N	2	14.767	82.758	22.900
11	CA	2	14.125	81.558	22.404
12	C	2	12.611	81.805	22.245
13	O	2	11.911	81.927	23.261
14	CB	2	14.358	80.407	23.367
15	N	3	12.194	81.901	20.988
16	CA	3	10.803	82.082	20.676
17	C	3	10.173	80.727	20.297
18	O	3	10.650	80.085	19.349
19	CB	3	10.652	83.058	19.522
20	N	4	9.165	80.348	21.074
21	CA	4	8.445	79.131	20.819
22	C	4	7.047	79.462	20.257
23	O	4	6.608	80.615	20.376
24	CB	4	8.305	78.330	22.102
25	N	5	6.442	78.450	19.647
26	CA	5	5.114	78.588	19.113
27	C	5	4.079	78.178	20.180
28	O	5	4.373	77.289	20.993
29	CB	5	4.955	77.714	17.881
30	N	6	2.945	78.866	20.145
31	CA	6	1.864	78.568	21.044
32	C	6	1.193	77.243	20.630
33	O	6	1.658	76.606	19.673
34	CB	6	0.841	79.690	21.018
35	N	7	0.165	76.881	21.388
36	CA	7	-0.594	75.695	21.099
37	C	7	-2.093	76.044	21.014
38	O	7	-2.691	76.384	22.046
39	CB	7	-0.369	74.657	22.184
40	N	8	-2.610	75.977	19.793
41	CA	8	-4.006	76.226	19.560
42	C	8	-4.854	75.414	20.559
43	O	8	-4.305	74.533	21.237
44	CB	8	-4.374	75.835	18.139
45	N	9	-6.130	75.774	20.624
46	CA	9	-7.058	75.079	21.473
47	C	9	-8.093	74.330	20.610

Table 16 continued

Atom Number	Atom type	Position in peptide	x	y	z
47	C	9	-8.036	74.312	20.948
48	O	9	-8.773	74.928	20.163
49	CB	9	-7.698	76.089	22.679
50	N	10	-8.021	72.999	21.144
51	CA	10	-8.966	72.137	20.488
52	C	10	-10.293	72.891	20.269
53	O	10	-10.367	73.727	19.356
54	CB	10	-9.213	70.899	21.332
55	N	11	-11.275	72.533	21.087
56	CA	11	-12.558	73.179	21.038
57	C	11	-13.561	72.288	20.278
58	O	11	-13.243	71.836	19.167
59	CB	11	-12.437	74.524	20.343
60	N	12	-14.696	72.054	20.925
61	CA	12	-15.750	71.281	20.326
62	C	12	0.000	-12.558	73.179
63	O	12	18.616	0.000	-12.558
64	CB	12	0.000	0.000	0.000

Table 17

Backbone 141					
Atom Number	Atom type	Position in peptide	x	y	z
0	N	0	0.000	0.000	0.000
1	CA	0	18.454	86.485	22.460
2	C	0	16.950	86.573	22.266
3	O	0	16.481	87.224	21.305
4	CB	0	0.000	0.000	0.000
5	N	1	16.227	85.893	23.151
6	CA	1	14.776	85.918	23.128
7	C	1	14.252	84.663	22.452
8	O	1	13.601	84.752	21.387
9	CB	1	14.299	87.132	22.349
10	N	2	14.573	83.520	23.055
11	CA	2	14.106	82.241	22.559
12	C	2	12.572	82.273	22.400
13	O	2	11.868	82.483	23.398
14	CB	2	14.499	81.135	23.523
15	N	3	12.141	82.099	21.156
16	CA	3	10.736	82.054	20.855
17	C	3	10.224	80.605	20.973
18	O	3	11.035	79.698	21.214
19	CB	3	10.489	82.573	19.449
20	N	4	8.911	80.468	20.833
21	CA	4	8.289	79.172	20.868
22	C	4	6.823	79.286	20.405
23	O	4	6.108	80.179	20.882
24	CB	4	8.338	78.611	22.279
25	N	5	6.465	78.404	19.478
26	CA	5	5.118	78.352	18.981
27	C	5	4.147	78.042	20.138
28	O	5	4.521	77.295	21.054
29	CB	5	4.999	77.280	17.911
30	N	6	2.972	78.656	20.055
31	CA	6	1.943	78.430	21.033
32	C	6	1.020	77.288	20.562
33	O	6	1.265	76.719	19.488
34	CB	6	1.130	79.697	21.234
35	N	7	0.034	76.991	21.401
36	CA	7	-0.938	75.983	21.081
37	C	7	-2.338	76.622	20.985
38	O	7	-2.577	77.649	21.637
39	CB	7	-0.939	74.903	22.150
40	N	8	-3.173	76.006	20.156
41	CA	8	-4.529	76.453	19.995
42	C	8	-5.492	75.437	20.641
43	O	8	-5.144	74.250	20.729
44	CB	8	-4.856	76.604	18.520
45	N	9	-6.629	75.957	21.087
46	CA	9	-7.649	75.129	21.670
47	C	9	-7.625	73.734	21.014

Table 17 continued

Atom Number	Atom type	Position in peptide	x	y	z
48	O	9	-6.531	73.205	20.765
49	CB	9	-9.013	75.766	21.470
50	N	10	-8.822	73.200	20.803
51	CA	10	-8.965	71.925	20.155
52	C	10	-10.460	71.616	19.939
53	O	10	-11.065	70.945	20.788
54	CB	10	-8.334	70.836	21.005
55	N	11	-10.983	72.148	18.840
56	CA	11	-12.353	71.910	18.476
57	C	11	-12.732	70.452	18.805
58	O	11	-12.400	69.551	18.020
59	CB	11	-12.548	72.168	16.992
60	N	12	-13.373	70.294	19.958
61	CA	12	-13.836	69.000	20.380
62	C	12	0.000	-12.353	71.910
63	O	12	18.541	0.000	-12.353
64	CB	12	0.000	0.000	0.000

Table 18

Backbone 144					
Atom Number	Atom type	Position in peptide	x	y	z
0	N	0	0.000	0.000	0.000
1	CA	0	18.480	86.428	22.392
2	C	0	16.967	86.551	22.343
3	O	0	16.431	87.361	21.553
4	CB	0	0.000	0.000	0.000
5	N	1	16.308	85.727	23.153
6	CA	1	14.861	85.759	23.256
7	C	1	14.262	84.643	22.416
8	O	1	13.512	84.919	21.454
9	CB	1	14.341	87.091	22.745
10	N	2	14.630	83.412	22.767
11	CA	2	14.106	82.241	22.093
12	C	2	12.565	82.287	22.092
13	O	2	11.968	82.501	23.158
14	CB	2	14.581	80.981	22.796
15	N	3	12.006	82.121	20.899
16	CA	3	10.578	82.090	20.743
17	C	3	10.094	80.628	20.667
18	O	3	10.880	79.754	20.273
19	CB	3	10.177	82.830	19.479
20	N	4	8.846	80.435	21.077
21	CA	4	8.236	79.135	21.020
22	C	4	6.879	79.228	20.292
23	O	4	6.338	80.337	20.167
24	CB	4	8.027	78.596	22.424
25	N	5	6.422	78.073	19.822
26	CA	5	5.148	77.990	19.162
27	C	5	4.052	77.645	20.190
28	O	5	4.068	76.532	20.737
29	CB	5	5.192	76.923	18.081
30	N	6	3.184	78.622	20.423
31	CA	6	2.076	78.436	21.319
32	C	6	1.134	77.348	20.765
33	O	6	1.402	76.819	19.676
34	CB	6	1.313	79.740	21.481
35	N	7	0.109	77.048	21.553
36	CA	7	-0.883	76.089	21.152
37	C	7	-2.256	76.780	21.027
38	O	7	-2.407	77.911	21.512
39	CB	7	-0.965	74.968	22.174
40	N	8	-3.167	76.084	20.357
41	CA	8	-4.509	76.574	20.198
42	C	8	-5.503	75.588	20.843
43	O	8	-5.193	74.391	20.931
44	CB	8	-4.832	76.735	18.722

Table 18 continued

Atom Number	Atom type	Position in peptide	x	y	z
45	N	9	-6.623	76.144	21.290
46	CA	9	-7.669	75.348	21.873
47	C	9	-8.201	74.343	20.832
48	O	9	-8.407	74.731	19.672
49	CB	9	-8.801	76.243	22.347
50	N	10	-8.360	73.106	21.286
51	CA	10	-8.894	72.067	20.448
52	C	10	-10.383	72.344	20.162
53	O	10	-11.124	72.681	21.097
54	CB	10	-8.745	70.719	21.133
55	N	11	-10.734	72.224	18.886
56	CA	11	-12.097	72.403	18.469
57	C	11	-12.907	71.126	18.774
58	O	11	-12.859	70.178	17.977
59	CB	11	-12.150	72.700	16.980
60	N	12	-13.575	71.155	19.921
61	CA	12	-14.414	70.059	20.322
62	C	12	0.000	-12.097	72.403
63	O	12	18.465	0.000	-12.097
64	CB	12	0.000	0.000	0.000

Example 4

The following method was used to identify high affinity binding peptides from Myelin Basic Protein (MBP). The binding affinities for a set of MBP peptides to HLA-DRB1*0401 have been experimentally determined and published. This set includes all possible 13 amino acid peptides from the MBP sequence which have a hydrophobic anchor residue at the P3 position. It is known that only such peptides bind to HLA-DR molecules with detectable affinity.

The same homology model of HLA-DRB1*0401 was used for this example as was used in Examples 1 and 2.

For each of the 13-mer peptides from the experimental determined set, a binding score was calculated as follows:

a) Calculate the steric overlap between the pocket bound peptide residue in the binding groove and an atom forming the pocket; this is value B.

b) Count the number of hydrogen bonds which could be formed between the pocket bound peptide residue and atoms forming the pocket; this is value C.

c) Calculate the strength of electrostatic interactions between any polar atoms of the pocket bound peptide residue and any polar atoms forming the pocket; this is value D.

d) Count the number of favourable contacts between the pocket bound peptide residue and atoms forming the pocket; this is value E.

e) These values were then transformed into a conformation score (Z) by using the following equation:

$$Z_n = cK_2C - cK_3D + cK_4E - cK_1B$$

Where K_1 to K_4 are constants and n is the sequence position of the peptide residue (numbered from 1 to the N-terminus to 13 at the C-terminus). K_1, K_2, K_3 and K_4 are equal to 100, 1500, 500 and 1000, respectively.

The conformation of each rotatable side-chain of the peptide residue was then altered by 15 degrees and the conformation score was recalculated.

The above steps were repeated for each residue of the peptide and the highest conformation score for each peptide residue was sued to determine the conformation score for the peptide.

At the point, the entire proceedings for establishing the conformation score for the peptide were repeated another 166 times, each time using a different peptide backbone form the library of peptide backbones.

The combination of peptide backbone and peptide side-chain conformations which gave the best conformation was then used to determine a binding score for the peptide.

The binding score was determined by establishing values of the following parameters:

a) Calculate the steric overlap between the pocket bound peptide residue in the binding groove and an atom forming the pocket; this is value B.

b) Count the number of hydrogen bonds which could be formed between the pocket bound peptide residue and atoms forming the pocket; this is value C.

c) Calculate the strength of electrostatic interactions between any polar atoms of the pocket bound peptide residue and any polar atoms forming the pocket; this is value D.

d) Count the number of favourable contacts between the pocket bound peptide residue and atoms forming the pocket; this is value E.

e) Calculate the hydrophobicity of the pocket bound peptide side chains using a hydrophobicity scale disclosed in Janin et al.

f) Calculate the number of MHC pocket residues which are paired with the pocket bound peptide residues. Pairing takes place if the centre of an atom from the MHC pocket residue and the centre of an atom from the pocket bound peptide residues are no more than the sum of their van der wall radii plus one Angstrom. The value A_n is calculated by summing the number of paired residues, where n is the number of the pocket. The values of A_n taking into account the pockets importance in binding are summed to give a value P.

The above values were then imported in to the following equation in order to determine the binding score (Y):

$$Y=P+bK_2C-bK_3D+bK_4E-bK_1B+bK_5He$$

Wherein the values bK_1, bK_2, bK_3, bK_4 and bK_5 are 2, 40, 600, 10 and 200 respectively.

As can be seen from the results in Table 19 the top four predicted scores pertain to four peptides which appear within the top five best binders.

Table 19

BB	PEPTIDE	AFFINITY	BINDING SCORE	D	E	F	B	P	Hc
104	HFFKNIVTPRTPP	40	4728	-0.12	11	17	97.7	3580	1.5
107	VHFFKNIVTPRTP	135	2125	-0.19	12	15	284.5	2255	0.2
104	PVVHFFKNIVTPR	161	4528	-0.06	15	12	337.6	4565	1.4
104	FSWGAEGQRPGFG	298	6205	-0.15	12	10	169.7	4670	-0.2
104	KGFKGVDAQGTLS	460	4353	-0.09	9	13	88.2	3145	1.9
112	KYLATASTMDHAR	479	2672	-0.09	13	15	106.8	1480	2.4
129	SKYLATASTMDHA	601	498	-0.06	11	13	275.7	620	0.4
141	RGLSLSRFSWGAE	1213	4140	-0.05	17	16	81.4	3455	1.7
62	TGILDSIGRFFGG	2942	337	0.04	21	17	25.3	-5	-0.6
0	RFFGGDRGAPKRG	3403	3218	-0.24	20	14	389.1	3100	1.6
104	NIVTPRTPPPSQG	6615	1971	0	10	11	305	2090	0.8
14	DSIGRFFGGDRGA	7268	1904	-0.08	8	15	37.3	1640	0.2
0	SRFSWGAEGQRPG	8352	1735	-0.08	20	13	466.8	1965	0.8
104	SKIFKLGGRDSRS	8494	1387	-0.1	10	10	149.2	825	2.8
118	SDYKSAHKGFKGV	8510	1864	-0.27	14	14	14.2	775	0.7
65	STMDHARHGFLPR	8860	1886	-0.21	14	15	191.3	1410	2.2
104	NPVVHFFKNIVTP	12870	1347	-0.11	12	10	332.5	1690	0.2
104	GTLSKIFKLGGRD	16000	4162	-0.11	17	10	118	3775	1.1
93	GRFFGGDRGAPKR	18467	244	-0.11	8	9	161	-175	2.3
75	KIFKLGGRDSRSG	25358	2185	-0.13	19	12	279.4	2060	1.4
0	FGYGGRASDYKSA	26397	1301	-0.12	16	16	306.1	1630	-0.4
0	PGFGYGGRASDYK	35200	3485	0.01	14	13	183.5	3165	1.4
144	GILDSIGRFFGGD	44400	2031	-0.09	21	14	32.1	1745	-0.5
134	KNIVTPRTPPPSQ	59000	1077	-0.04	9	10	45.9	340	3.1
0	KGVDAQGTLSKIF	100000	2067	-0.11	24	15	695.2	2795	0.3

KEY - BB = NUMBER OF THE BACKBONE CHOSEN FROM THE LIBRARY

CLAIMS

1. A method for the prediction of the binding affinity of a peptide to a major histocompatibility (MHC) class II molecules comprising;

a) ascertaining the characteristics of a MHC molecule binding groove,

b) presenting a selected peptide to the MHC molecule and ascertaining a first conformation score for each pocket bound peptide side-chain,

c) amending the conformation of each pocket bound peptide side-chain and ascertaining a second conformation score,

d) repeating step 3 with alternative conformations of each peptide pocket bound side-chain,

e) choosing the highest conformation score for each pocket bound peptide side-chain in each binding groove pockets, herein known as `the pocket', and

f) combining the highest conformation score for each pocket and ascertaining a binding score for the complete peptide.

2. A method according to claim 1 which further comprises the step of compiling information on all peptide fragments in a protein and comparing the binding scores.

3. A method according to any preceding claim wherein the conformation score is ascertained by at least one of the following parameters:

a) the number of favourable contacts between MHC residues forming one of the pockets and the pocket bound peptide residue; this is value E

b) the steric overlap between the pocket bound peptide residue bound in the pocket and an atom forming the pocket; this is value B,

c) the number of hydrogen bonds which could be formed between the pocket bound peptide residue and an atom forming the pocket; this is value C,

d) the strength of electrostatic interactions between any

polar atoms of the pocket bound peptide residue and any polar atoms forming the pocket; this is value D.

4. A method according to claim 3 wherein the steric overlap between the pocket bound peptide residue and the atoms forming the pocket can not be greater than 0.35 Angstroms.

5. A method according to claim 3 wherein a favourable contact occurs when an atom from an MHC residue and an atom from the peptide residue have their centres separated by no more than the sum of their radii plus 0.5 Angstroms and are not overlapping.

6. A method according to the preceding claims wherein values B to E are imported into a first equation, to give a conformation score (Z)

7. A method according to claim 6 wherein the first equation is $Z_n=(cK_2C)-(cK_3D)+(cK_4E)-(cK_1B)$, where cK_1 to cK_4 are constants and n is the number of the pocket.

8. A method according to claim 7 wherein cK_1 is between 50 and 150.

9. A method according to claim 7 wherein cK_2 is between 1000 and 2000.

10. A method according to claim 7 wherein cK_3 is between 250 and 750.

11. A method according to claim 7 wherein cK_4 is between 500 and 1500.

12. A method according to any preceding wherein the Z_n value for a pocket is multiplied by a coefficient, L, depending on the pockets importance in binding, to give a second Z_n value.

CLAIMS

1. A method for the prediction of the binding affinity of a peptide to a major histocompatibility (MHC) class II molecules comprising;

a) ascertaining the characteristics of a MHC molecule binding groove,

b) presenting a selected peptide to the MHC molecule and ascertaining a first conformation score for each pocket bound peptide side-chain,

c) amending the conformation of each pocket bound peptide side-chain and ascertaining a second conformation score,

d) repeating step 3 with alternative conformations of each peptide pocket bound side-chain,

e) choosing the highest conformation score for each pocket bound peptide side-chain in each binding groove pockets, herein known as `the pocket', and

f) combining the highest conformation score for each pocket and ascertaining a binding score for the complete peptide.

2. A method according to claim 1 which further comprises the step of compiling information on all peptide fragments in a protein and comparing the binding scores.

3. A method according to any preceding claim wherein the conformation score is ascertained by at least one of the following parameters:

a) the number of favourable contacts between MHC residues forming one of the pockets and the pocket bound peptide residue; this is value E

b) the steric overlap between the pocket bound peptide residue bound in the pocket and an atom forming the pocket; this is value B,

c) the number of hydrogen bonds which could be formed between the pocket bound peptide residue and an atom forming the pocket; this is value C,

d) the strength of electrostatic interactions between any

polar atoms of the pocket bound peptide residue and any polar atoms forming the pocket; this is value D.

4. A method according to claim 3 wherein the steric overlap between the pocket bound peptide residue and the atoms forming the pocket can not be greater than 0.35 Angstroms.

5. A method according to claim 3 wherein a favourable contact occurs when an atom from an MHC residue and an atom from the peptide residue have their centres separated by no more than the sum of their radii plus 0.5 Angstroms and are not overlapping.

6. A method according to the preceding claims wherein values B to E are imported into a first equation, to give a conformation score (Z)

7. A method according to claim 6 wherein the first equation is $Z_n=(cK_2C)-(cK_3D)+(cK_4E)-(cK_1B)$, where cK_1 to cK_4 are constants and n is the number of the pocket.

8. A method according to claim 7 wherein cK_1 is between 50 and 150.

9. A method according to claim 7 wherein cK_2 is between 1000 and 2000.

10. A method according to claim 7 wherein cK_3 is between 250 and 750.

11. A method according to claim 7 wherein cK_4 is between 500 and 1500.

12. A method according to any preceding wherein the Z_n value for a pocket is multiplied by a coefficient, L, depending on the pockets importance in binding, to give a second Z_n value.

13. A method according to any of the preceding claims wherein all the Z values are summed to give a value J.

14. A method according to any of the preceding claims wherein the MHC residue is paired with the pocket-bound peptide residue if an atom from the MHC residue and an atom from the pocket-bound peptide residue have their centres separated by no more than the sum of their van der Waal radii plus one Angstrom.

15. A method according to claim 14 wherein a value A_n is calculated by summing the pairwise interaction frequencies of paired residues.

16. A method according to either claim 14 or 15 wherein the value A_n for a pocket is multiplied by a coefficient, X, depending on the pockets importance in binding.

17. A method according to claim 16 wherein the A_n value for the pockets are summed to give a value P.

18. A method according to any preceding claim wherein the binding score is ascertained by at least one of the following parameters

a) the number of groove-bound hydrophobic residues; this is value F,

b) the number of non groove-bound hydrophilic residues; this is value G,

c) the number of peptide residues deemed to fit within their respective binding pocket; this is value H.

19. A method according to any one of claims 13 to 18 wherein values F, G, H, J and P are imported into a second equation to give a first binding score, Y.

20. A method according to claim 19 wherein the second algorithm is $Y=J*F^2*(G*H+1)+ P$.

21. A method according to claim 1-17 wherein the hydrophobicity of the pocket bound peptide side chains is evaluated using a hydrophobicity scale; this is value He.

22. A method according to claim 21 wherein the hydrophobicity scale ranges from -1.8 for lysine to 0.9 for cysteine.

23. A method according to either of claims 21 or 22 wherein $Y=(bK_2C)-(bK_3\ D)+(bK_4E)-(bK_1B)+(bK_5He)+P$.

24. A method according to claim 23 wherein bK_1 is between 1 and 5.

25. A method according to claim 23 wherein bK_2 is between 20 and 60.

26. A method according to claim 23 wherein bK_3 is between 300 and 900.

27. A method according to claim 23 wherein bK_4 is between 1 and 20.

28. A method according to claim 23 wherein bK_5 is between 1 and 800.

29. A method according to any preceding claim wherein the steps in claim 3 are repeated for each pocket and each conformation of the peptide residue in said pocket.

30. A method according to claim 29 wherein the conformation of the peptide is altered by rotating a side chain of the peptide residue by a pre-determined amount.

31. A method according to either claim 29 or 30 where in the conformation of the peptide is altered by changing the conformation of the peptide backbone.

32. A method according to any preceding claim wherein the steps are repeated using different peptides from a protein.

33. A method according to any of the preceding claim wherein the binding scores (Y) for different peptides are tabulated and compared.

34. A method according to any of the preceding claim which is used in the manufacture of a vaccine derived from a peptide identified by said method.

35. A method according to any of the preceding claims which is used to remove potentially immunogenic sequences from a protein and thus reduce said proteins immunogenicity when administered to an organism.

36. A computer conditioned to receive information characterising a peptide bound to the MHC molecule and to utilise said information to perform a procedure having the following steps;

a) ascertaining the characteristics of a MHC molecule binding groove;

b) presenting a selected peptide, which is selected by a predetermined program, to the MHC molecule and ascertaining a first conformation score;

c) amending the conformation of the peptide, by way of a predetermined program, and ascertaining a second conformation score;

d) repeating step 3 with other conformations of the peptide;

e) selecting the peptide conformation with the highest conformation score; and

f) calculating the binding score from the conformation score.

37. A computer according to claim 36 further comprising a step (7) which comprises repeating steps 1-4 with other peptide fragments in the protein

to generate information on all peptide fragments in a protein

so that a comparison can be made of the strength of the binding between the peptide and the MHC molecule.

38. A computer according to either claim 36 or 37 further comprising a step (8) which comprises altering the conformation of the backbone of the peptide fragment.

39. A pharmaceutical composition produced resultant upon to a method as claimed in anyone of claims 1 to 35.

FIG. 1

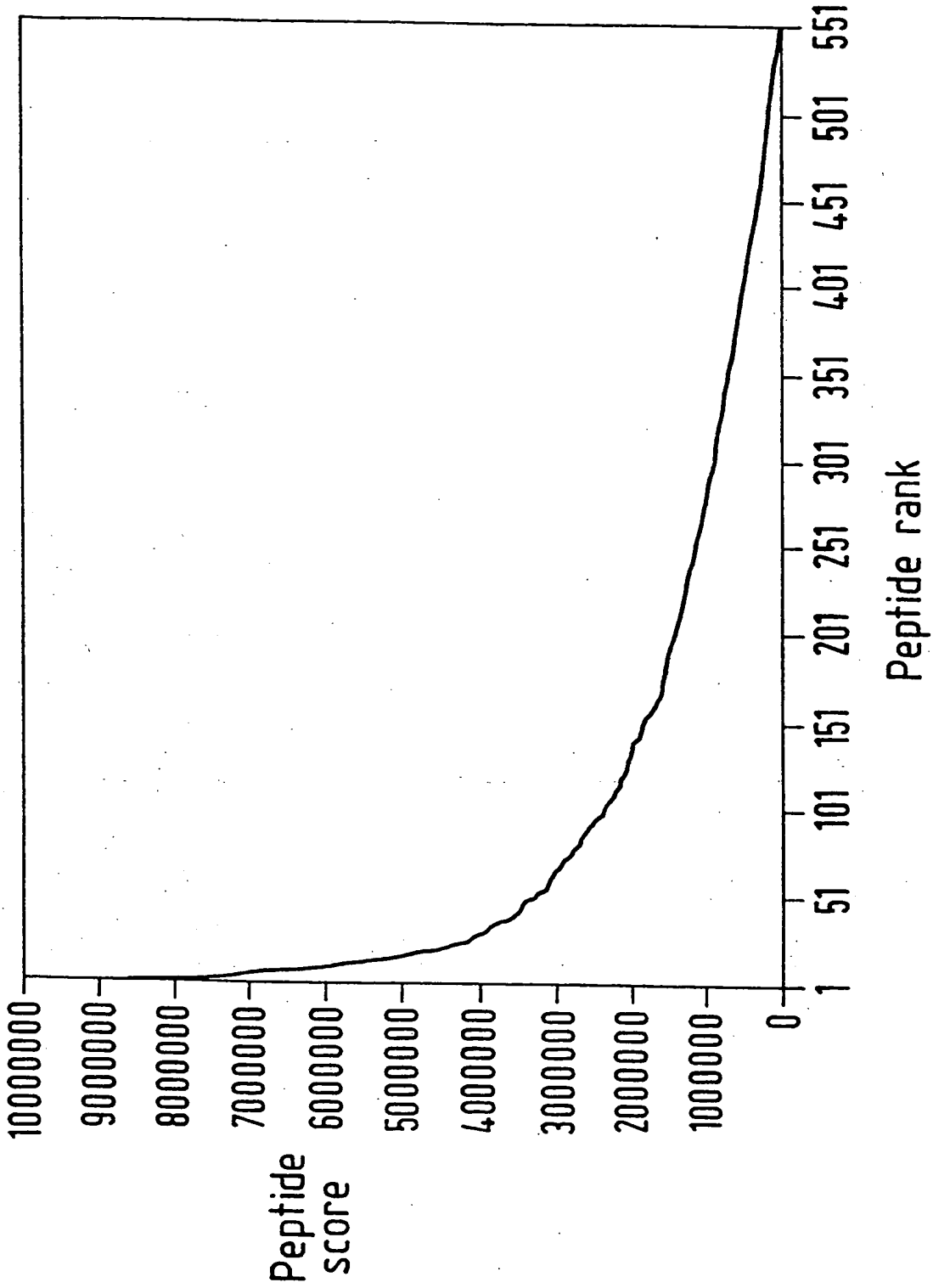
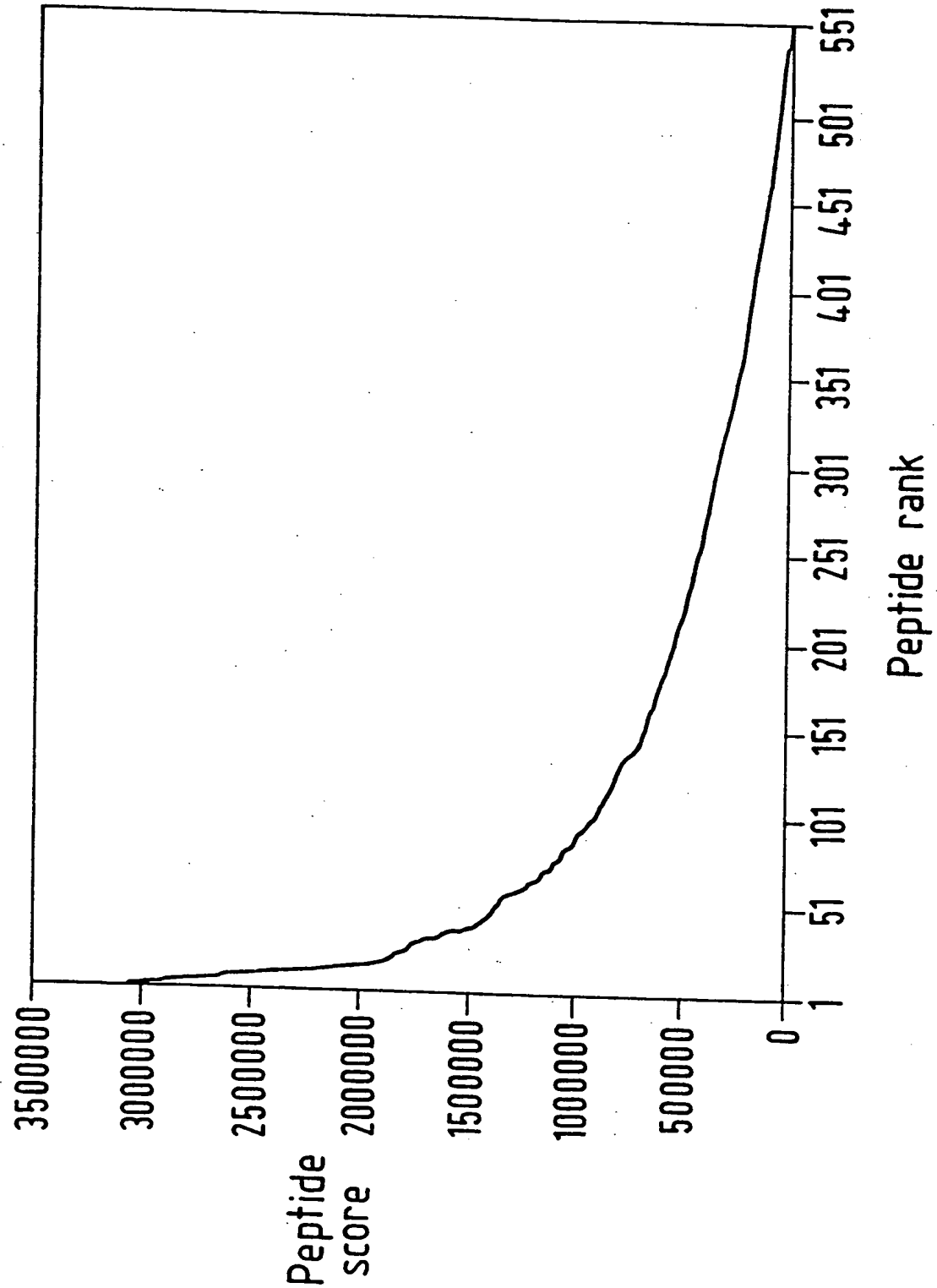


FIG. 2



INTERNATIONAL SEARCH REPORT

International Application No
PCT/[redacted]8/01801

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 G01N33/569 G01N33/564 G01N33/566 C07K14/705

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 G01N C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 95 31483 A (ECLAGEN LTD) 23 November 1995 see page 2, line 23 - line 28 see page 5, line 5 - line 12	1-35
X	---	39
X,P	WO 97 40852 A (ANERGEN INC) 6 November 1997 see claims 31,32	39
A,P	---	1-35
	-/--	

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
22 October 1998	05/11/1998

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Van Bohemen, C

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

International Application No
PCT/[redacted]98/01801

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
T	T.E. JOHANSEN ET AL.: "Peptide binding to MHC class I is determined by individual pockets in the binding groove." SCANDINAVIAN JOURNAL OF IMMUNOLOGY, vol. 46, no. 2, 1 August 1997, pages 137-146, XP002081826 oxford uk see the whole document	1-35,39

2

INTERNATIONAL SEARCH REPORT

International application No.
PCT/GB 98/01801

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. [X] Claims Nos.: 36-38
because they relate to subject matter not required to be searched by this Authority, namely:
Rule 39.1(i) PCT - Mathematical method

2. [ ] Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

3. [ ] Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. [ ] As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. [ ] As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. [ ] As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. [ ] No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

[ ] The additional search fees were accompanied by the applicant's protest.

[ ] No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/98/01801

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9531483 A	23-11-1995	AU 2452195 A	05-12-1995
		CA 2190101 A	23-11-1995
		EP 0759944 A	05-03-1997
		JP 10500670 T	20-01-1998
WO 9740852 A	06-11-1997	AU 2421397 A	19-11-1997